# WRKY transcription factors involved in salicylic acid-induced defense gene expression

Verk, M.C. van

# 5

# WRKY Transcription Factors Involved in Activation of SA Biosynthesis Genes

Marcel C. van Verk, John F. Bol, and Huub J.M. Linthorst

Manuscript in Preparation

## ABSTRACT

Increased defense against a variety of pathogens in plants is achieved through activation of a mechanism known as systemic acquired resistance (SAR). The broad-spectrum resistance brought about by SAR is mediated through salicylic acid (SA). An important step in SA biosynthesis in Arabidopsis is the conversion of chorismate to isochorismate through the action of isochorismate synthase, encoded by the *ICS1* gene. Also $AVR_{PPHB}$ *SUSCEPTIBLE 3* (*PBS3*) plays an important role in SA metabolism, as *pbs3* mutants accumulate drastically reduced levels of SA-glucoside, a putative storage form of SA. Bioinformatics analysis previously performed in our group identified WRKY28 and WRKY46 as possible regulators of *ICS1* and *PBS3*. Expression studies with *ICS1 promoter::β-glucuronidase* (*GUS*) genes in *Arabidopsis thaliana* protoplasts cotransfected with *35S::WRKY28* showed that over expression of WRKY28 resulted in a strong increase in GUS expression. Moreover, qRT-PCR analyses indicated that the endogenous *ICS1* and *PBS3* genes were highly expressed in protoplasts overexpressing WRKY28 or WRKY46, respectively. Electrophoretic shift assays indentified three potential WRKY28 binding sites in the *ICS1* promoter, positioned -445, -460 and -121 base pairs upstream of the transcription start site. Mutation of these sites in protoplast transactivation assays showed that the binding sites at -445 and -460 are functionally important for activation of the *ICS1* promoter. Chromatin immunoprecipitation assays with haemagglutinin-epitope tagged WRKY28 showed that the region of the *ICS1* promoter containing the binding sites at -445 and -460 was highly enriched in the immunoprecipitated DNA.




## INTRODUCTION

Because of their sessile nature, plants have evolved very sophisticated mechanisms to actively cope with different sorts of stresses. The various defense mechanisms that can be initiated are controlled by signaling molecules like salicylic acid (SA) or jasmonic acid (JA) or by combinations of these signal compounds. SA accumulates locally in infected leaves, as well as in non-infected systemic leaves after infection with biotrophic pathogens and mediates the induced expression of defense genes, resulting in an enhanced state of defense known as systemic acquired resistance (SAR) (Métraux *et al.*, 1990; Malamy *et al.*, 1990; Dempsey *et al.*, 1999; Ryals *et al.*, 1996; Glazebrook 2005). SAR is a long-lasting broad spectrum resistance against a variety of pathogenic fungi, bacteria and viruses (Thomma *et al.*, 2001; Durrant and Dong 2004). Also exogenous application of SA results in induced expression of defense related genes (White, 1979; van Loon *et al.*, 1997). Among the genes that are induced during SAR is a set of genes collectively known as PR (pathogenesis-related) genes, with members encoding anti-fungal ß-1,3-glucanases (PR-2), chitinases (PR-3, PR-4) and PR-1, which are often used as molecular markers for SAR (Hunt *et al.*, 1996; van Loon *et al.*, 1997; Mou *et al.*, 2003; Durrant and Dong 2004).

Biosynthesis of SA can occur via two different pathways, the pathway that synthesizes SA from phenylalanine (Lee *et al.*, 1995), and the isochorismate pathway. Inhibition of the phenylalanine pathway still allows accumulation of SA (Yalpani *et al.*, 1993; Mauch-Mani *et al.*, 1996). An important step in the isochorismate pathway is the conversion of chorismate to isochorismate (ICS). Expression of a bacterial *ICS* gene in plants causes accumulation of SA, constitutive expression of *PR* genes and constitutive SAR (Verberne *et al.*, 2000), whereas the *sid2* mutant corresponding with a defective *ICS1* gene, is compromised in accumulation of SA and unable to mount SAR (Wildermuth *et al.*, 2001; Nawrath and Métraux, 1999). Expression of the *ICS1* gene is rapidly induced after infection (Wildermuth *et al.,* 2001). $AVR_{PPHB}$ *SUSCEPTIBLE 3* (*PBS3*), of which the pathogen-induced expression is highly correlated with expression of *ICS1*, is acting downstream of SA. In the *pbs3* mutant, accumulation of SA-glucoside and expression of *PR-1* are drastically reduced. The *PBS3* gene product is a member of the auxin-responsive GH3 family of acyl-adenylate/thioester forming enzymes of which some have been shown to catalyze hormone–amino acid conjugation, like the protein encoded by the *JAR1* gene catalyzes the formation of JA-isoleucine. Although the observation that PBS3 is not active on SA, INA and Chorisamate leads to the hypothesis that PBS3 must be placed upstream of SA. (Jagadeeswaran *et al.*, 2007; Nobuta *et al.*, 2007; Okrent *et al.*, 2009).

After perception of pathogen attack by cytoplasmic TIR-NB-LRR receptors, several genes are involved in initiation of the defense response upstream of *ICS1*. One of these genes is *ENHANCED DISEASE SUSCEPTIBILITY 1* (*EDS1*), which is probably activated after elicitor perception (Wirthmueller *et al.*, 2007). EDS1 heterodimerizes with *PHYTOALEXIN DEFICIENT 4* (*PAD4*) and their nuclear localization is important for subsequent steps in the signaling pathway (Aarts *et al.*, 1998; Feys *et al.*, 2001). Both *EDS1* and *PAD4* are induced by pathogen infection and SA application. The accumulation of SA is also regulated by $Ca^{2+}$ via EDS1. *EDS1* expression is repressed by the $Ca^{2+}$/calmodulin-binding transcription factor Serine/threonine protein kinase 1 (AtSR1) that can bind to the *EDS1* promoter and repress *EDS1* gene expression (Du *et al.*, 2009). Another enhanced disease susceptibility gene (*EDS5*) that is also situated upstream of SA biosynthesis is expressed at high levels upon pathogen infection in an *EDS1*- and *PAD4*-dependent manner (Rogers and Ausubel, 1997). The *eds5* mutant plants are no longer able to accumulate high levels of SA upon pathogen infection and are unable to initiate the SAR response (Nawrath and Métraux, 1999).

Although many mutants have been reported to affect SA accumulation, no direct transcriptional regulators of genes like *ICS1* or *PBS3* have been identified. For *ICS1* the presence of many TGAC core sequences, as present in the binding sites for WRKY transcription factors, has been hypothesized to be important for transcriptional regulation of *ICS1* gene expression (Eulgem and Somssich, 2007). Here we describe two WRKY transcription factors that were previously identified in our group via a bioinformatics analysis to be closely co-expressed with *ICS1* and *PBS3*. Co-expression analyses in protoplasts showed that WRKY28 and WRKY46 positively regulated the expression of *ICS1* and *PBS3*, respectively. In addition, the binding sites for WRKY28 in the *ICS1* promoter were identified.

Our results indicate that WRKY28 and WRKY46, which themselves are both rapidly induced by pathogen elicitors (Navarro *et al.*, 2004; He *et al.*, 2006), link pathogen-triggered defense gene expression to the accumulation of SA via induction of *ICS1* and *PBS3* gene expression.

## RESULTS

### WRKY28 Activates *ICS1::GUS* Gene Expression in Arabidopsis Protoplasts

The co-expression analysis from Chapter 4 indicated that WRKY28 and WRKY46 could play a role in regulation of *ICS1* and *PBS3*. To verify if WRKY28 and WRKY46 can act as positive transcriptional regulators of *ICS1* and/or *PBS3* gene expression we performed transactivation assays in Arabidopsis protoplasts. Protoplasts were cotransfected with plasmids containing either the *WRKY28* or *WRKY46* coding region behind the *35S* promoter, together with a plasmid containing the GUS reporter gene cloned behind the 1kb promoter region of *ICS1* or of *PBS3*. As controls, the *promoter::GUS* fusions were cotransfected with an "empty" plasmid lacking the *WRKY28* or *WRKY46* coding region. The results of these transactivation assays are shown in Figure 1. *ICS1* promoter-directed GUS expression is increased approximately 4-fold by WRKY28 in comparison to the empty vector control. No increase is observed after cotransfection with the WRKY46 plasmid. In the case of *PBS3* promoter-directed GUS expression, neither WRKY28 nor WRKY46 positively stimulated gene expression.

To analyze the effect of WRKY28 and WRKY46 on expression of endogenous *ICS1* and *PBS3* genes, Arabidopsis protoplasts were transfected with *35S::WRKY28* or *35S::WRKY46* plasmids and incubated overnight, after which total RNA was isolated for qRT-PCR analysis of the expression of the endogenous *ICS1* and *PBS3* genes. Often, WRKYs positively regulate their own expression (Pandey *et al.*, 2009) and therefore expression of the endogenous *WRKY28* and *WRKY46* genes was also investigated. The constitutive housekeeping genes *Actin3*, *Actin7*, *Actin8* and β-*Tubelin* were used as controls. The results of the qRT-PCR analyses are shown in Figure 2. WRKY28 overexpression resulted in a 4.5 fold increase of *ICS1* mRNA. This confirms the presence of WRKY28 responsive elements in the *ICS1* promoter, at least part of which are present in the 1 kb fragment analyzed in Figure 1. WRKY28 did not increase expression of the *PBS3* gene. Apparently neither the 1 kb fragment of the *PBS3* promoter (Fig. 1) nor the full-length promoter contains WRKY28 responsive elements. Overexpression of WRKY46 had no effect on expression of the *ICS1* gene, indicating that the full-length promoter of this gene does not contain WRKY46 responsive elements. However, WRKY46 overexpression resulted in a 4-fold increase of the *PBS3* mRNA level. This indicates that the *PBS3* promoter contains WRKY46 responsive elements, located more than 1 kb upstream of the transcription start site. Obviously, there is no positive effect of WRKY28 or WRKY46 on the expression of the corresponding endogenous WRKY genes, but both WRKYs did have a slightly negative effect on the expression of the endogenous *WRKY28* gene.

### Characterization of the WRKY28 Binding Sites in the *ICS1* Promoter

As a first step towards the characterization of WRKY28 binding sites in the *ICS1*

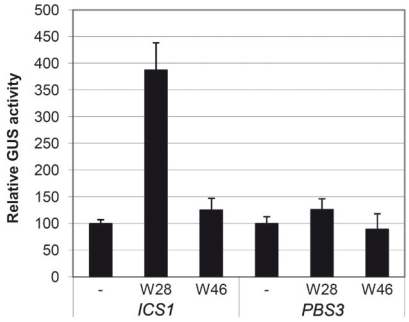


Figure 1. Transactivation of *ICS1::GUS* and *PBS3::GUS* promoter fusions by WRKY28 and WRKY46 in Arabidopsis protoplasts. The fusions contained promoter sequences of 1 kb upstream of the transcription start sites of the *ICS1* or *PBS3* genes. Protoplasts were transfected with vector pRT101 containing *35S::WRKY28* (W28) or *35S::WRKY46* (W46) inserts, or with the empty vector (minus sign). In the left three columns, the protoplasts were co-transfected with the *ICS1::GUS* fusion, in the right three columns, the protoplasts were co-transfected with the *PBS3::GUS* fusion. The columns represent the average relative GUS expression observed in four experiments. GUS expression induced in the presence of the empty pRT101 vector was taken as 100%. Error bars represent the SEM.



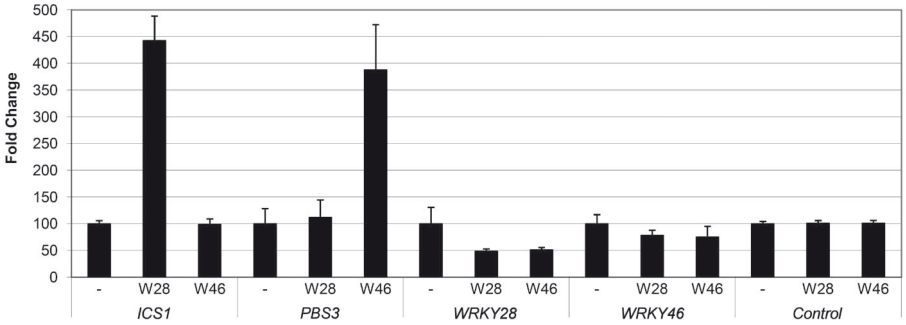


Figure 2. Effect of WRKY28 and WRKY46 on the expression of several endogenous Arabidopsis genes. Expression of ICS1, PBS3, WRKY28, WRKY46 and four household genes in Arabidopsis protoplasts was measured by qRT-PCR. Expression of each gene was measured in protoplasts transfected with the empty pRT101 vector (minus sign) or with the pRT101 vector containing *35S::WRKY28* (W28) or *35S::WRKY46* (W46) inserts. Bars represent the average fold change in mRNA levels observed in three experiments. mRNA levels in protoplasts transfected with the empty pRT101 vector were taken as 100%. The control represents the average of the data obtained with the four household genes. Error bars represent the SEM.

promoter, a region of this promoter of 960 base pairs (bp) upstream of the transcription start site was divided by PCR into six overlapping fragments. After labeling, the fragments were assayed for their ability to bind to purified glutathione S-transferase (GST)/WRKY28 fusion protein expressed in *E. coli* using electrophoretic mobility shift assays (EMSAs). The results

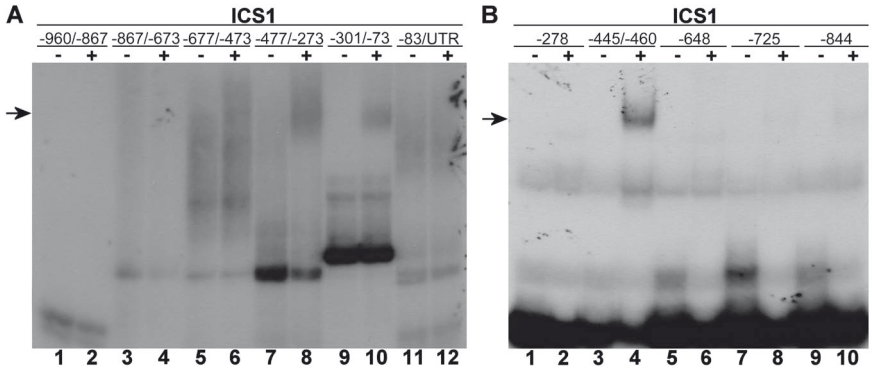


Figure 3. Binding of WRKY28 to *ICS1* promoter fragments. (A) EMSAs were done with six overlapping promoter fragments spanning the 960 bp sequence upstream of the transcription start site of the ICS1 gene. The borders of these fragments are given on top of the lanes. (B) EMSAs were done with promoter fragments of 30 bp, each containing a TGAC core sequence (positions -278, -445/-460, -648, -725) or a WK-like box (-844) in the center. The location of these sequences in the *ICS1* promoter is given on top of the lanes. The promoter fragments were incubated with recombinant GST/WRKY28 fusion protein (plus-signs) or without this protein (minus-signs). The position of protein-DNA complexes is indicated by an arrow.



are shown in Figure 3A. It is evident that in the presence of GST/WRKY28, part of the probes corresponding to fragments -477/-273 (Fig 3A, compare Lanes 7 and 8) and -301/-73 (Fig. 3A, compare Lanes 9 and 10) shifted to a higher position in the gel. No such band shifts were observed with the other promoter fragments.

WRKY proteins are generally considered to bind to the consensus W-box sequence TTGAC(C/T) (Eulgem *et al.*, 2000). The 1 kb *ICS1* promoter does not contain a true W-box, although a number of TGAC core sequences is present (positions -725, -648, -460, -445 and -278). Furthermore, a WK-like box (TTTTCCA) that resembles the WK-box TTTTCCAC identified by van Verk *et al.* (2008) is present at position -844. In addition to the six promoter fragments that spanned the 960 bp sequence, we prepared 30-bp promoter fragments that contained a TGAC core sequence or the WK-like box in the center. (The two inverted TGAC sequences at positions -445 and -460 were present in one 30-bp fragment.) The results of EMSAs with these fragments as probes are shown in Figure 3B. The shifted band in Lane 4 indicates that the 30-bp fragment containing the two cores at -445 and -460 was bound to GST/WRKY28 protein and this could explain the observed shift in the -477/-273 fragment shown in Figure 3A, Lane 8. With none of the other WK-like or W-box core sequences a shift could be observed (Fig. 3B, Lanes 2, 6, 8, 10). To verify the binding specificity of the 30-bp fragment containing the TGAC cores at positions -445 and -460, competition experiments were done with 50- and 250-fold excess unlabelled fragments (Fig. 4B). Evidently, addition of a 250-fold excess unlabelled fragment completely outcompeted the binding to the probe (Fig 4B, Lane 4), indicating that this *ICS1* promoter fragments specifically interacted with WRKY28.

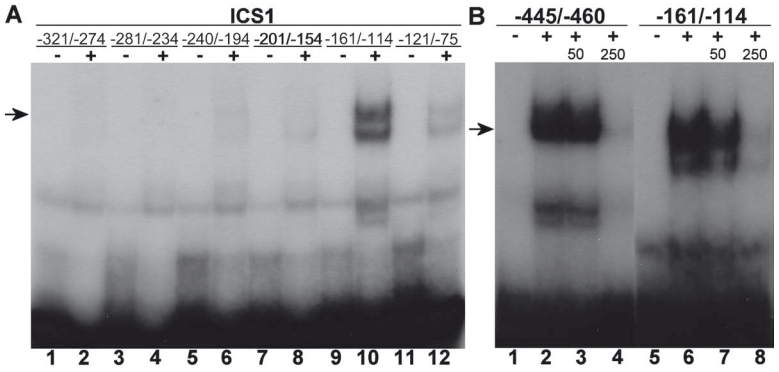


Figure 4. Binding of WRKY28 to *ICS1* promoter fragments. (A) EMSAs were done with *ICS1* promoter fragments corresponding to the positions given on top of the lanes. (B) EMSAs were done with a 30-bp fragment of the *ICS1* promoter containing TGAC core sequences at position -445 and -460 (four left lanes) and with a fragment of the *ICS1* promoter from position -161 to -114 (four right lanes. The EMSAs in panel B were done without addition of unlabeled competitor DNA, or in the presence of a 50-fold or 250-fold excess of unlabeled competitor DNA as indicated on top of the lanes. The promoter fragments were incubated with recombinant GST/WRKY28 fusion protein (plus-signs) or without this protein (minus-signs). The position of protein-DNA complexes is indicated by an arrow.



We speculated that the two TGAC core sequences at -445 and -460 could be binding sites for WRKY28 and set out to further investigate which site is responsible for the observed shift. To determine if the TGAC cores at -445 and -460 are involved in binding WRKY28, a scanning analysis was performed with a series of annealed complementary oligonucleotide probes in which the core sequences were changed to CCGG (Fig. 5C, m1, m2 and m1+2). The results of EMSAs with these fragments are shown in Figure 5A, Lanes 1 to 8. Mutation of either the core at -460 (m1) or at -445 (m2) does not abolish binding of WRKY28 to the fragment (Fig. 5A, compare Lanes 2, 4 and 6). However, mutation of both cores in mutant m1+2 disrupts binding (Fig. 5A, Lane 8). This suggests that both binding sites are equally important.

To further analyze the requirements for binding of WRKY28, pairwise mutations of the sequence around the core at -445 were scanned in an m1 background (Fig. 5C). The results are shown in Figure 5A, Lanes 9 to 24. Mutation of m2.1 and m2.4 show binding to WRKY28 (Fig. 5A, Lanes 10 and 16). As would be expected, mutations within the core sequence completely abolished binding of WRKY28 (m2.2 and m2.3, Fig. 5A, Lanes 12 and 14). Since the TGAC core at -460 has TC upstream of the core and the inverted core at -445 has a CT in this position, we checked to which extend the T or C nucleotides are important for binding. Changing CT to TC resulted in a binding of WRKY28 that was as strong as to the wild type sequence (m2.5, Fig. 5A, Lane 18). Changing CT to TT significantly lowered binding (m2.6, Fig. 5A, lane 20), suggesting that the presence of a C at either position -1 or -2 from the core is important for binding WRKY28. We further analyzed the effect of mutations at positions -3/-4 and +3/+4 from the core. Pairwise mutation of nucleotides at

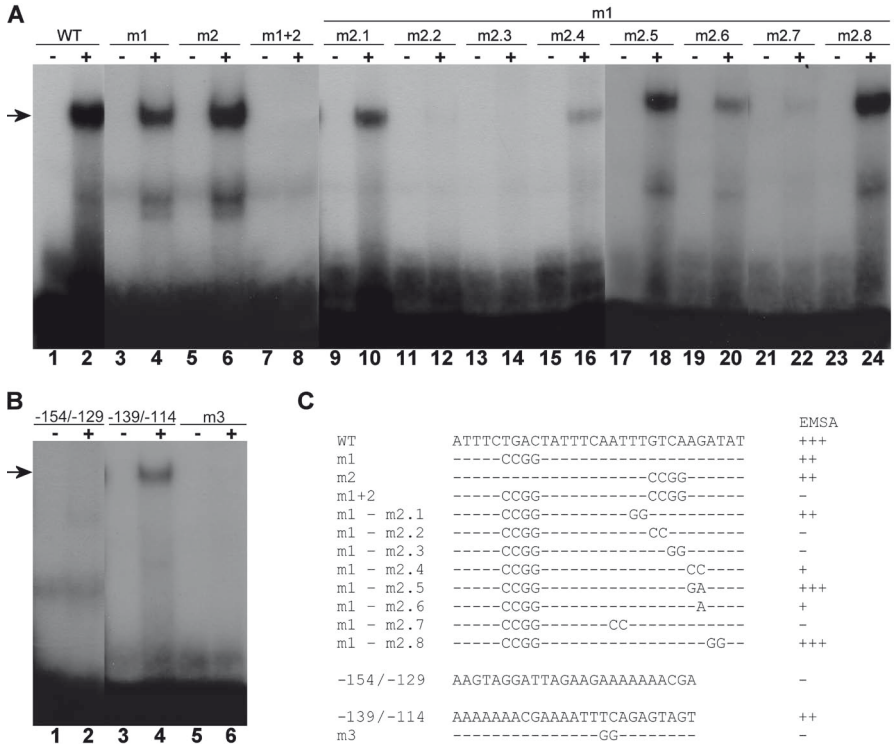


Figure 5. Binding of WRKY28 to mutated *ICS1* promoter fragments. EMSAs were done with annealed 30-bp oligonucleotides containing the *ICS1* promoter region indicated -445/-460 in the legend of Figure 3 with mutations as indicated in panel C (A). Annealed 25-bp oligonucleotides corresponding to promoter regions indicated in panel C, bottom two lines (B). Plus signs above the lanes indicate binding mixtures containing 0.5 µg recombinant GST/WRKY28. Minus signs above the lanes indicate binding mixtures without recombinant protein. The position of the protein-DNA complexes is indicated by an arrow. Plus and minus signs in panel C indicate the relative abundance of the shifted probe.

-3/-4 did not alter the binding of WRKY28 (m2.8, Fig. 5A, Lane 24), however no shift was observed when the nucleotides at +3/+4 were mutated, indicating that this flanking sequence is important for binding of WRKY28 (m2.7, Fig. 5A, Lane 22).

To more directly demonstrate that the binding sites at positions -460 and -445 are involved in WRKY28 activation of *ICS1* gene expression, mutants m1, m2 and m1+2 were introduced in the 1 kb *ICS1* promoter and their effects studied in cotransfection experiments in Arabidopsis protoplasts. While cotransfection of *35S::WRKY28* with the wild-type *ICS1 promoter::GUS* increased GUS expression approximately 3.5-fold over the background level, expression dropped significantly with promoter constructs containing the m1 or m2 mutation (Fig. 6). Combination of m1 and m2 (m1+2) did not lower GUS expression more than the single mutations (Fig. 6). This result supports the notion that WRKY28 activates *ICS1* expression through specific binding sites in the promoter at

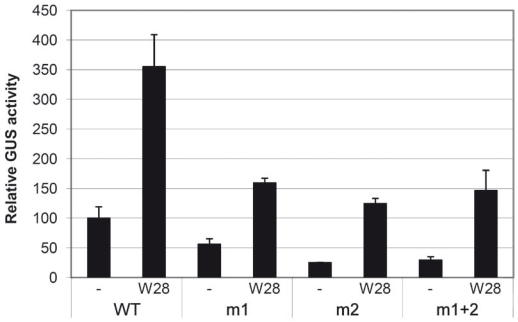


Figure 6. Transactivation of *ICS1::GUS* genes with mutations in WRKY28 binding sites. Protoplasts were transfected with 2 µg of wild-type *promoter::GUS* constructs or *promoter::GUS* constructs containing the mutations m1, m2 or m1+2 as indicated in Figure 5C. W28, cotransfection with 6 µg of expression vector pRT101 containing *35S::WRKY28*. Minus signs, cotransfection with 6 µg of empty expression vector. The bars represent the percentage of GUS activity from triple experiments relative to that of the protoplasts cotransfected with the *promoter::GUS* construct and an empty expression vector, which was set to 100%. Error bars represent the SEM.



-445 and -460 bp upstream of the transcription start site.

The EMSA experiments shown in Fig. 3A indicated that promoter fragment -301/-73 contains a binding site for WRKY28 (Fig. 3A, lane 10). Probably, this site is not located in the 30 bp sequence centered around the TGAC core sequence at position -278, because no WRKY28 binding site was detectable in the 30 bp fragment corresponding to this position (Fig. 3B, lane 2). Outside this 30 bp sequence, no sequences with similarity to known WRKY binding sites were detectable in the -301/-73 fragment. To delineate WRKY28 binding sites in the -301/-73 sequenceof the *ICS1* promoter, the entire region was divided into six overlapping fragments of approximately 50 bp. These fragments, offered as annealed complementary oligonucleotides in EMSAs, were evaluated for binding to GST/WRKY28. The results of this experiment are shown in Figure 4A. With fragments within the regions -321 to -154 and -121 to -75 bp upstream of the transcription start site no shift or only a very faint shift was observed (Fig. 4A, Lanes 1-8, and 11-12). However, a prominent shift occurred with fragment -161/-114 (Fig. 4A, Lane 10), suggesting that this region contains a WRKY28 binding site. Competition with 50- and 250-fold excess unlabeled fragment -161/-114 confirmed the specificity of the fragment for binding WRKY28 (Fig. 4B, right panel).

The region between -161 and -114 lacks a TGAC core sequence. To identify the WRKY28 binding site in this region, the left half (nucleotides -154 to -129) and the right half (nucleotides -139 to -114) were separately assayed for binding WRKY28 in EMSA (Fig. 5B, Lanes 1-4). Only with the right half a shift could be observed (Fig. 5B, Lane 4). This region harbors the sequence TTCA (-121), which in reverse orientation is somewhat similar to the W-box core TGAC. To investigate if this sequence is part of the binding site in this region, the middle two nucleotides were mutated (Fig. 5C). This mutation abolished the ability of WRKY28 to bind (m3, Fig. 5B, Lane 6), suggesting that

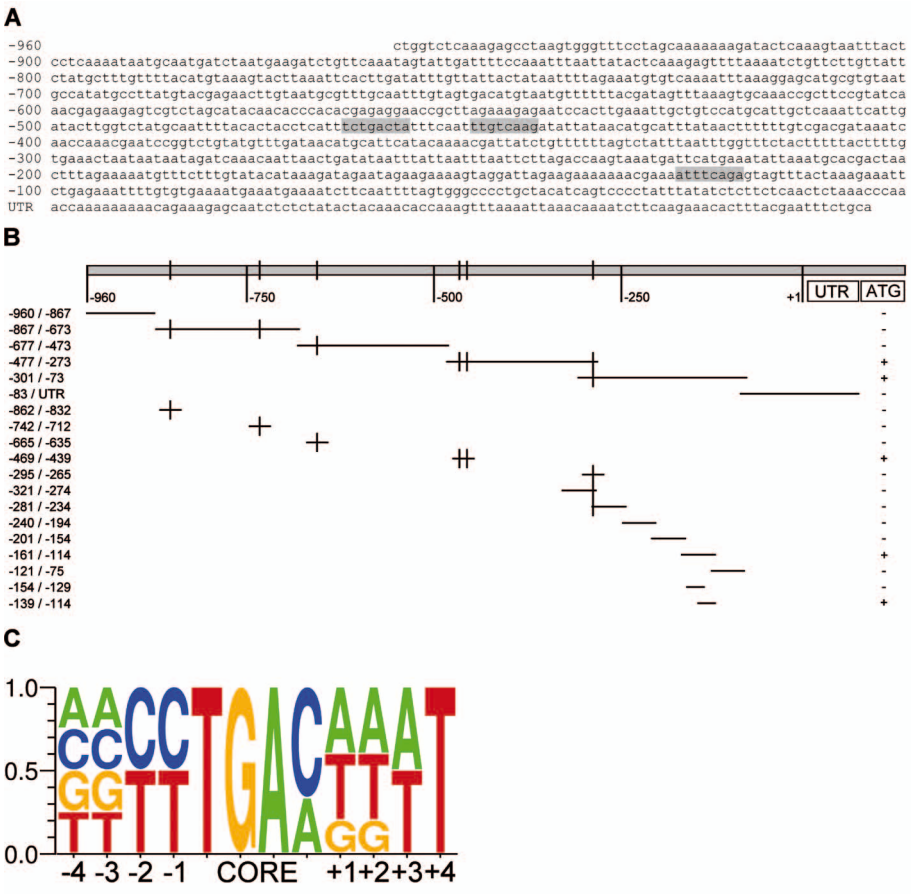


Figure 7, Summary of Electrophoretic Mobility Shift Assays. Sequence of the 960 bp *ICS1* promoter. The indentified WRKY28 binding sites are indicated against a grey background (A). Schematic representation of the *ICS1* promoter fragments analyzed by EMSA (B). Consensus WRKY28 binding sequence deduced from the EMSAs (C). In B, plus-signs in the right column indicate fragments that produced band shifts; minus-signs, fragments that did not produced a band shift. The position of the WK-like sequence or TGAC core sequences is indicated by vertical lines.

besides W-boxes with the conserved TGAC core, variants with cores like TGAA may also facilitate binding to WRKY transcription factors.

To summarize the results of the EMSAs, Figure 7A shows the 960 bp of the *ICS1* promoter with the characterized WRKY28 binding sites indicated against a grey background. A schematic representation of the fragments tested in EMSAs for binding WRKY28 is given in Figure 7B. The consensus binding sequence generated using the program WebLogo (Crooks *et al.*, 2004) by combination of the characterized binding sites and the results of the mutational analysis of the binding site at -445, is shown in Figure 7C.

**Chromatin Immunoprecipitation Analysis of WRKY28**

The transactivation experiments in protoplasts and the *in vitro* binding studies described above support a role for WRKY28 as a transcriptional activator of *ICS1*. To check if WRKY28 is able to bind to the *ICS1* promoter *in vivo*, chromatin immunoprecipitation (ChIP) assays were set up using Arabidopsis protoplasts, as described by Lee *et al.* (2007). The *WRKY28* coding sequence was fused to a haemagglutinin (HA) tag and expressed in Arabidopsis protoplasts. The resulting WRKY28-HA fusion protein was able to induce GUS expression when cotransfected with an *ICS1 promoter::GUS* construct, indicating that the HA tag did not interfere with WRKY28's functionality (Results not shown).

For ChIP analysis WRKY28-HA or unfused HA were expressed in protoplasts. After 24h incubation, chromatin complexes were cross-linked using formaldehyde. Upon shearing by sonication, the fragmented chromatin was incubated with monoclonal anti-HA antibodies overnight, after which immunoprecipitated complexes were captured using magnetic protein G beads and extensively washed. DNA eluted from the beads was analyzed by qPCR with primers corresponding to six overlapping regions of the *ICS1* promoter (Fig. 8A). qPCRs with primers corresponding to the coding region of *PR1* and the promoter region of *PDF1.2* were included as controls. The results are shown in Figure 8B. With the primer sets corresponding to *PR1* and *PDF1.2* no specific products were amplified, indicating that these sequences were absent from the immunoprecipitated chromatin. While no specific PCR products were amplified with primer sets A, B, D and F, it is evident that the region corresponding to the *ICS1* promoter bordered by primers C was highly enriched in the immunoprecipitated chromatin from the WRKY28-HA transfected protoplasts (38-fold in comparison to the control). This region contains the two WRKY28 binding sites at -445 and -460 as determined by EMSA (Fig. 5A). A similar result was obtained with a primer pair covering a smaller region containing the two binding sites (Results not shown). Surprisingly, approximately 20-fold enrichment was observed when primer set E was used with the chromatin precipitate from the WRKY28-HA transfected protoplasts (Fig. 8B). This primer set encompasses a further upstream region of the *ICS1* promoter, which was not found to bind to WRKY28 in the EMSA assays. However, while the amplification efficiencies of the qPCRs with all other primer sets was always above 70%, the qPCRs with primers E had a low efficiency of less than 30%, suggesting amplification of non-specific DNA sequences. Moreover, the atypical melting curve of the PCR product obtained with primers E indicated that the product was heterogeneous. Indeed, when checked on gel, primer set C produced a discrete band of the expected size, while the product of primers E consisted of a mixture of differently sized fragments (Results not shown). In conclusion, the ChIP assays indicated that WRKY28 specifically binds to the *ICS1* promoter *in vivo*, most probably to one or both binding sites at position -460 and -445 upstream of the transcription start site.

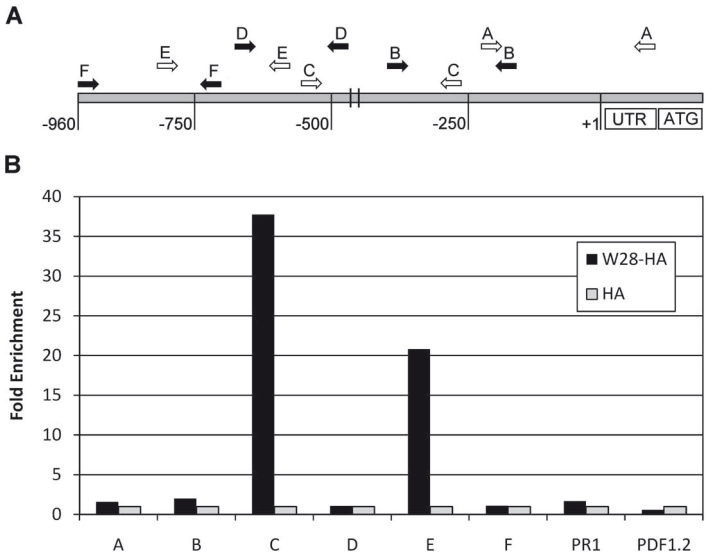


Figure 8. Chromatin Immunoprecipitation assay. Schematic representation of the location of primers corresponding to regions of the *ICS1* gene used in the ChIP assays (A). Fold enrichment of immunoprecipitated DNA from protoplasts expressing WRKY28-HA versus protoplasts expressing unfused HA (B). The position of the WRKY28 binding sites at -445 and -460 is indicated.



## DISCUSSION

### WRKY28 and WRKY46 Activate Expression of *ICS1* and *PBS3*, Respectively

Our *in silico* co-expression analysis of Arabidopsis transcription factor genes and genes involved in stress signaling suggested many putative new components of the signal transduction pathways (Chapter 4). Among the genes resulting from this screening were two encoding WRKY transcription factors linked to genes involved in SA metabolism. The gene encoding the type II member WRKY28 was found to be closely co-regulated with the *ICS1* gene involved in SA biosynthesis, whereas the type III *WRKY46* gene linked to *PBS3*. Based on this finding we decided to investigate the effects of these WRKYs on transcriptional activation of *ICS1* and *PBS3*. Indeed, overexpression of WRKY28 in Arabidopsis protoplasts lead to enhanced GUS activity from a co-expressed *GUS* reporter gene under control of a 1 kb *ICS1* promoter, and also expression of the endogenous *ICS1* gene was increased (Figs. 1 and 2). Likewise, overexpression of WRKY46 resulted in increased accumulation of *PBS3* mRNA, supporting the notion that WRKY46 is a transcriptional activator of *PBS3* (Fig. 2). GUS activity was not enhanced from a co-expressed 1 kb *PBS3* promoter*::GUS* gene. This indicates that WRKY46 activates the *PBS3* gene by binding at a position in the promoter further upstream than

1 kb. However, we cannot exclude the possibility that the 1 kb promoter used for the construction of the reporter construct and which was derived from curated genome sequence data by The Arabidopsis Information Resource (TAIR), is not the actual *PBS3* promoter. A detailed analysis of the region upstream of the coding sequence in the Arabidopsis genome shows that the intron of almost 1 kb suggested to be present in the 5'-UTR contains several putative binding sites for transcription factors like WRKYs and TGAs. It will be interesting to investigate if the suggested "intron" is the actual *PBS3* promoter.

**DNA Binding Site of WRKY28**

Several studies on DNA binding characteristics of WRKY transcription factors have led to the generally accepted consensus binding sequence TTGAC[C/T], commonly referred to as the W-box (Rushton *et al.*, 1996; de Pater *et al.*, 1996; Wang *et al.*, 1998; Eulgem *et al.*, 2000; Chen and Chen, 2000; Cormack *et al.*, 2002; Eulgem and Somssich, 2007; Ciolkowski *et al.*, 2008). Recently, we identified a variant binding site for the tobacco NtWRKY12 transcription factor (van Verk *et al.*, 2008). NtWRKY12 binds to a WK-box (TTTTCCAC), which deviates significantly from the W-box consensus sequence. Based on this finding we have suggested that the TTTTCCA sequence in EMSA probes binding to the barley transcription factor SUSIBA2 could be this WRKY's WK-like binding site (van Verk *et al.*, 2008; Sun *et al.*, 2003).



In this study we have characterized three sites in the *ICS1* promoter that have a high affinity for WRKY28. The consensus WRKY28 binding site that emerged from this analysis has some characteristics that differ from the W-box consensus (Fig. 7). We found that, unlike the consensus W-box, a C may be present at position -1 in front of the TGAC core, and although a T is also allowed at -1, a C is then required at -2. Similarly, for the sequence after the core, in two of the binding sites an A is present at +1, which in the W-box is usually either a C or a T. Remarkably, one of the WRKY28 binding sites has TGAA, instead of TGAC as core. These findings indicate that the consensus W-box is not the only WRKY binding site.

To disable binding of WRKY28 to the 30-bp EMSA probe harboring the binding sites at -460 and -445, mutation of both these sites was necessary. With only one site intact, binding was still possible (Fig. 5A, Lanes 4 and 6). Nevertheless, with the 1 kb promoter, mutation of only one of the sites had a severe effect on reporter gene expression and expression was not further reduced when both sites were mutated. Apparently, for transcriptional activation both sites are required. Possibly, activation requires that WRKY28 binds as a dimer, similar to WRKYs 18, 40 and 60, which were found to form functionally relevant homo- and heterodimers (Xu *et al.*, 2006).

The transactivation experiments also showed that mutation of the sites at -460 (m1) and -445 (m2) did not completely knock out reporter gene expression. In comparison to the GUS activity obtained with the wild type construct, approximately 20% remained. Furthermore, the reduction in basal expression levels seen with the mutant *ICS1* promoters in the absence of overexpressed WRKY28 indicates that also endogenous factors binding to the sites at -460 and -445 contribute

to the expression level. qRT-PCR has shown that the WRKY28 gene is much higher expressed in protoplasts than in suspension cells from which the protoplasts were made (Results not shown), suggesting that possibly these factors include endogenous WRKY28. Moreover, the residual GUS expression remaining with the m1, m2 and m1+2 mutant promoters could indicate that other sites in the ICS1 promoter are still able to bind WRKY28. Further analyses are required to see if the binding site identified in promoter fragment -139/-114 is a candidate for such sites.

### Integrated Model for Regulation of SA Biosynthesis by WRKY28 and WRKY46

The combined results of the work described here, lead us to propose the following model for the induction of SA biosynthesis upon pathogen attack. Induction of the basal defense response starts with the detection of a pathogen-associated molecular pattern (PAMP), like in the case of flagellin, which is perceived by the FLS receptor. The activated FLS receptor triggers a MAP kinase cascade (MAPKKK/MEKK1?, MKK4/5, MPK3/6), which leads to transcriptional activation of the *WRKY28* gene (Navarro *et al.*, 2004). Transcription factor WRKY28 subsequently activates expression of the *ICS1* gene, through binding the promoter at the two binding sites at -460 and -445 and possibly at other sites, resulting in synthesis of ICS that catalyzes SA production. How the activated MAP kinase induces *WRKY28* gene expression remains a matter of speculation. The activated MAPK could activate an as of yet unknown transcription factor on stand by or release one from a repressor complex, or it may function itself as activator of *WRKY28* expression.

Less is known about the role of the product of the *PBS3* gene. It is rapidly induced in plants recognizing pathogens carrying virulence factors, like in the case of *Pseudomonas syringae* containing AVR4 (He *et al.,* 2006). A function in SA metabolism has been suggested based on its effect on SA-glucoside accumulation and its similarity to phytohormone-amino acylases (Nobuta *et al.*, 2007; Jagadeeswaran *et al.*, 2007). *PBS3* gene expression is repressed by high levels of SA, indicating that it is more likely that PBS3 functions early in the defense response before SA levels start to rise (Okrent *et al.*, 2009). Similarly, WRKY46 expression is rapidly induced upon infection and our finding that it enhances *PBS3* gene expression suggests an early role in R-gene-mediated defense. Figure 9 shows the placement of the two WRKYs in the SA-signaling pathways.

## MATERIALS AND METHODS

### Protoplast Preparation, Transfection and Analysis

For transactivation and qRT-PCR experiments, protoplasts were prepared from cell suspensions of *Arabidopsis thaliana* ecotype Col-0 according to van Verk *et al.*, (2008).

For transactivation experiments protoplasts were co-transfected with 2 µg of plasmid carrying *ICS1 promoter::GUS* (bp -1 to -960 relative to the transcriptional start site) or, *PBS3 promoter::GUS* (bp -1 to -1kb relative to the transcriptional start site) construct and 6 µg of *35S::effector* plasmid pRT101. As a control, cotransfection of promoter*::GUS* construct with the empty expression vector pRT101 was carried out. The protoplasts were harvested 16 hrs after transformation and GUS activity was determined (van der Fits and Memelink, 1997). GUS activities from triplicate experiments were

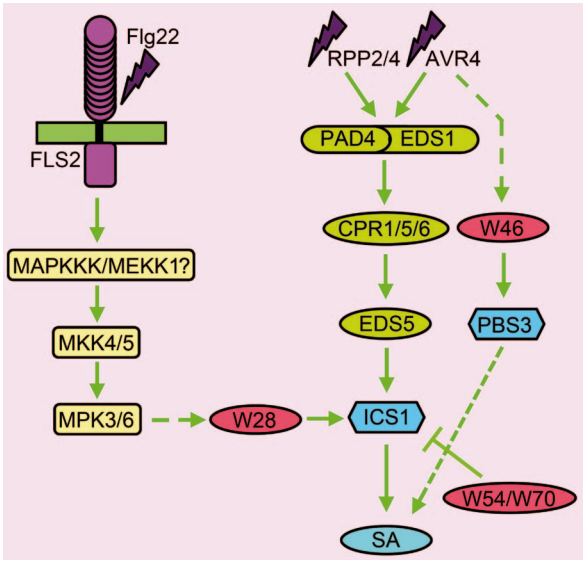


Figure 9. Model for regulation of SA biosynthesis by WRKY28 and WRKY46. Upon infection with a pathogen expressing flagellin (Flg22) or avirulence genes (RPP2/4 or AVR4), *WRKY28* or *WRKY46* are rapidly induced. Activation of FLS2 receptor by Flg22 results in activation of a MAPK cascade, which leads to induction of *WRKY28* expression, which subsequently activates *ICS1* gene expression leading to SA production. Avirulence factors like AVR4 trigger SA production through a pathway involving genes *PAD4*, *EDS1*, *CPR1/5/6*, *EDS5* and *ICS1*. WRKY46 is rapidly synthesized and either directly or indirectly positively regulates *PBS3* gene expression, having a positive influence on SA metabolism.



normalized against total protein level.

To analyze effects on expression of endogenous genes by WRKY28 and WRKY46, protoplasts were transfected with 6 µg of *35S::WRKY28* or *35S::WRKY46* expression plasmids. After 24h protoplasts were harvested and total RNA isolated. Total RNA was treated with DNAse using the Turbo DNA-free kit (Ambion) and cDNA was synthesized using the universal first strand cDNA synthesis kit (Fermentas). Expression of endogenous genes was determined by qPCR using primers listed in Table 1. qPCR was performed using a standard Phusion high fidelity polymyerase reaction (Finzymes) supplemented with 0.145µl Tween-20, 1.45µl Glycerol, 1 mM $MgCl_2$ and 1x Sybr green (Roche #70140720) per 50µl reaction and analyzed using a BioRad Chromo4 qPCR machine.

### Electrophorectic Shift Assays

Protein for EMSAs was purified from *E. coli* transformed with pGEX-KG constructs containing the open reading frame of *WRKY28* in frame behind the GST open reading frame, according to van Verk *et al.*, (2008).

EMSAs were performed essentially as described by Green *et al.* (1989). DNA probes for the EMSA assays were obtained by slowly cooling down mixtures of equimolar amounts of complementary oligonucleotides with a 5'-GGG overhangs from 95°C to room temperature or by PCR of the *ICS1* promoter fragments. Annealed oligonucleotides were subsequently end-filled and PCR fragments labeled throughout using Klenow fragment and [α-$^{32}$P]-dCTP, after which unincorporated label was removed by Autoseq G-50 column chromatography (Amersham-Pharmacia Biotech).

Table 1, Oligonucleotides used for qRT-PCR and ChIP qPCR analysis

| | | |
|---|---|---|
| qPCR-*Actin 3* | F | 5'-CCTCATGCCATCCTCCGTCT-3' |
| | R | 5'-CAGCGATACCTGAGAACATAGTGG-3' |
| qPCR-*Actin 7* | F | 5'-AGTGGTCGTACAACCGGTATTGT-3' |
| | R | 5'-GAGGAAGAGCATACCCCTCGTA-3' |
| qPCR-*Actin 8* | F | 5'-AGTGGTCGTACAACCGGTATTGT-3' |
| | R | 5'-GAGGATAGCATGTGGAAGTGAGAA-3' |
| qPCR-*β-Tubelin* | F | 5'-GGAAGAAGCTGAGTACGAGCA-3' |
| | R | 5'-GCAACTGGAAGTTGAGGTGTT-3' |
| qPCR-*ICS1* | F | 5'-GGAACAGTGTCATCTGATCGTAATC-3' |
| | R | 5'-CATTAAACTCAACCTGAGGGACTG-3' |
| qPCR-*PBS3* | F | 5'-CGTACCGATCGTGTCATATGAAG-3' |
| | R | 5'-CTTCACATGCTTGGTTATAACTTGC-3' |
| qPCR-*WRKY28* | F | 5'-CAAGAGCCTTGATCGATCATTG-3' |
| | R | 5'-GCAAGCCCAACTGTCTCATTC-3' |
| qPCR-*WRKY46* | F | 5'-CATGAGATTGAGAACGGTGTG-3' |
| | R | 5'-CTGCCATTAAGAGAGAGACATTACATTC-3' |
| ChIP-A | F | 5'-GTCAAAGCTTGCACGACTAACTTTAGAAAAATG-3' |
| | R | 5'-CAGTGGATCCTGCAGAAATTCGTAAAGTGTTTC-3' |
| ChIP-B | F | 5'-GTCAAAGCTTCAACCAAACGAATCCGGTCTGT-3' |
| | R | 5'-GAAGAGATCTATTTCATTTTCACACAAAATTTCTC-3' |
| ChIP-C | F | 5'-GTCAAAGCTTCAAACGAGAAGAGTCGTCTAGC-3' |
| | R | 5'-GGGTCAGTTAATTGTTTGATCTATTATTATTAG-3' |
| ChIP-D | F | 5'-GTCAAAGCTTGCCATATGCCTTATGTACGAGA-3' |
| | R | 5'-AGAAAGATCTTAGTGTAAAATTGCATAGACCAAG-3' |
| ChIP-E | F | 5'-GTCAAAGCTTCTATGCTTTGTTTTACATGTAAAG-3' |
| | R | 5'-GGGAAAAACATTACATGTCACTACAAATTGCAA-3' |
| ChIP-F | F | 5'-GTCAAAGCTTCTGGTCTCAAAGAGCCTAAGTG-3' |
| | R | 5'-GGGCTCCTTTAAATTTTGACACATTTCTAAAAT-3' |
| ChIP-*PR1* | F | 5'-GTTCTTCCCTCGAAAGCTCAAGAT-3' |
| | R | 5'-CACCTCACTTTGGCACATCCG-3' |
| ChIP-*PDF1.2* | F | 5'-TATACTTGTGTAACTATGGCTTGG-3' |
| | R | 5'-TGTTGATGGCTGGTTTCTCC-3' |

EMSA reaction mixtures contained 0.5 µg purified protein, 3 µL 5x gel shift binding buffer [20% glycerol, 5 mM MgCl2, 2.5 mM EDTA, 2.5 mM DTT, 250 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.25 mg $mL^{-1}$ poly(dI-dC) x poly(dIdC) (Promega)] in a total volume of 14 µL. After 10-min incubation at room temperature, 1 µL containing 30,000 cpm of labeled probe, representing approximately 0.01 pmol, was added and incubation was continued for 20 min at room temperature. 50- and 250-fold molar excess of unlabelled competitor was added for some reactions, representing 0.50 and 2.50 pmol respectively. The total mixture was loaded onto a 5% polyacrylamide gel in Tris-borate buffer and electrophoresed. After electrophoresis, the gel was dried, autoradiographed, and analyzed using X-ray film.

**Chromatin Immunoprecipitation**

For ChIP assays, protoplasts were prepared as described above and transfected with 6 µg of *35S::WRKY28-HA* or *35S::HA* constructs in plasmid pRT101. After 24h, protoplasts were harvested and ChIP assays were conducted as described by Lee *et al.* (2007), with minor modifications. After formaldehyde fixation, the chromatin of the protoplasts was isolated and extensively sheared by

sonication to obtain feagment sizes between 300-400 bp. Rat anti-HA monoclonal antibodies (clone 3F10, Roche) and Dynabeads Protein G magnetic beads (Invitrogen) were used to immunoprecipitate the genomic fragments. qPCRs were performed on the immunoprecipitated DNA using primers corresponding to six overlapping regions of the *ICS1* promoter as shown in Figure 8A. qPCRs with primers specific for the coding region of the *PR1* gene and the promoter of *PDF1.2* gene of Arabidopsis were used as controls. The primers used for the ChIP assays are listed in Table 1.

**ACKNOWLEDGEMENTS**

We gratefully acknowledge the group of dr. Ji Hoon Ahn for providing a detailed protocol for ChIP analysis in Arabidopsis protoplasts. Rob van Eck and Roy Baas are acknowledged for their help with some of the experiments and we are grateful to Ward de Winter for technical assistance.